Basic Inferential Data Analysis Instructions: Comparing tooth growth by dose and supp

Jae Kwon

March 3, 2017

Overview

Analysis of the ToothGrowth data in the R datasets package. Some basic exploratory data analyses are performed, and confidence interval and hypothesis test were done to compare tooth growth by supp and dose.

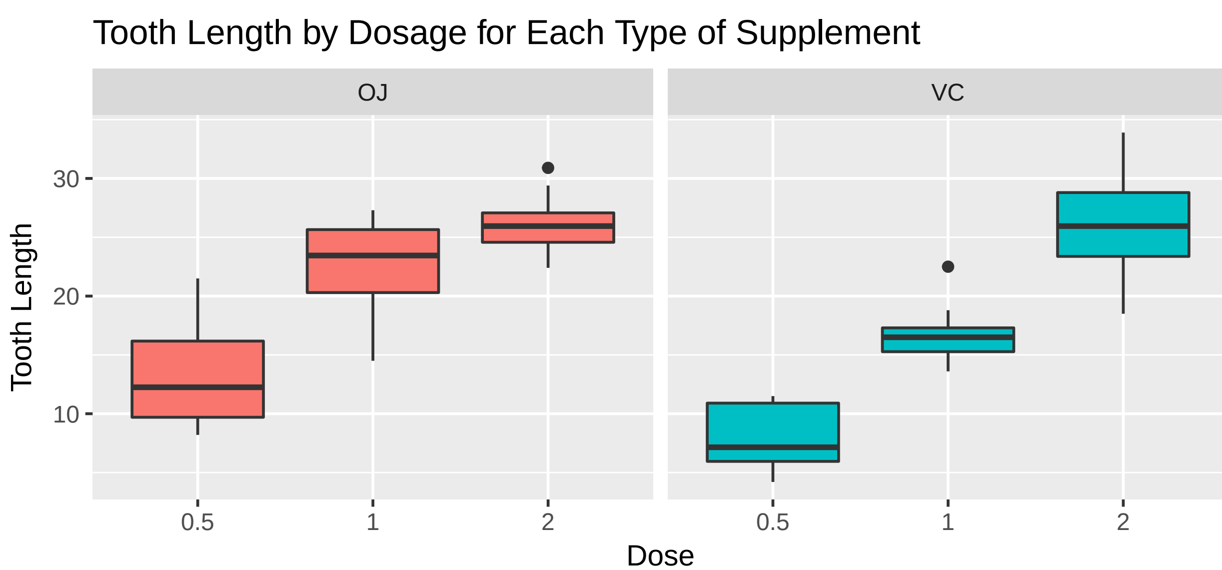
Basic Summary of Data

Load the data and see a quick summary of it:

```
data("ToothGrowth")
tg <- ToothGrowth
```

```
library(ggplot2)

ggplot(tg, aes(x = factor(dose), y = len, fill = supp)) +
       facet_grid(.~supp) + geom_boxplot(show_guide=FALSE) +
       xlab('Dose') + ylab('Tooth Length') +
       ggtitle('Tooth Length by Dosage for Each Type of Supplement')
```



Confidence Intervals & Hypothesis Test

By Supplement Type:

Hypothesis Test #1:

Null hypothesis: The two supplements result same tooth growth disregard the dosage.

```
# by supplement type
hyp_test <- t.test(len ~ supp, data =tg)
c.int <- hyp_test$conf.int
pval <- hyp_test$p.value
```

Confidence Interval with significance value of 0.05:

```
c.int
```

```
## [1] -0.1710156  7.5710156
## attr(,"conf.level")
## [1] 0.95
```

p-value is:

```
pval
```

```
## [1] 0.06063451
```

Since the p-value of 0.061 is larger than the significance value of 0.05, we fail to reject the null hypothesis. There is no strong evidence that the length of tooth is affected by the type of supplement. The average difference in tooth length across the two supplement types is close to 0.

If we assume that the difference is 0, it falls into the confidence interval.

By Dose Level:

Hypothesis Test #2:

Null hypothesis: For the dosage of 0.5 mg/day, the two supplements result same tooth growth.

```
htest1 <- t.test(len ~ supp, data = subset(tg, dose == 0.5))
```

Confidence Interval:

```
c.int1 <- htest1$conf.int
c.int1
```

```
## [1] 1.719057 8.780943
## attr(,"conf.level")
## [1] 0.95
```

P-value:

```
pval1 <- htest1$p.value
pval1
```

```
## [1] 0.006358607
```

Since P-value is less than the significant level of 0.05, we reject the null hypothesis. In addition, the confidence interval does not include 0 that there is a difference in tooth growth length. OJ with dosage of 0.5 mg/day has greater effect on the tooth growth.

Hypothesis Test #3:

Null hypothesis: For the dosage of 1.0 mg/day, the two supplements result same tooth groowth.

```
htest2 <- t.test(len ~ supp, data = subset(tg, dose == 1))
```

Confidence Interval:

```
c.int2 <- htest2$conf.int
c.int2
```

```
## [1] 2.802148 9.057852
## attr(,"conf.level")
## [1] 0.95
```

P-value:

```
pval2 <- htest2$p.value
pval2
```

```
## [1] 0.001038376
```

Since P-value is less than the significant level of 0.05, we reject the null hypothesis. In addition, the confidence interval does not include 0 that there is a difference in tooth growth length. OJ with dosage of 1.0 mg/day has greater effect on the tooth growth.

Hypothesis Test #4:

Null hypothesis: For the dosage of 2.0 mg/day, the two supplements result same tooth groowth.

```
htest3 <- t.test(len ~ supp, data = subset(tg, dose == 2.0))
```

Confidence Interval:

```
c.int3 <- htest3$conf.int
c.int3
```

```
## [1] -3.79807  3.63807
## attr(,"conf.level")
## [1] 0.95
```

P-value:

```
pval3 <- htest3$p.value
pval3
```

```
## [1] 0.9638516
```

Since P-value is much greater than the significant level of 0.05, we cannot reject the null hypothesis. The confidence interval does include 0 that it is significant there the two supplements at dosage of 2.0 mg/day have no different effect on tooth growth.

Conclusion & Assumptions:

Assumptions:

- The length of the tooth is normally distributed.
- There are no other factors that affect the length of the tooth.

Oj has a greater effect on tooth growth at dosage of 0.5 mg/day and 1.0 mg/day. However, at dosage of 2.0 mg/day, the two supplements does not show difference in tooth growth. Overall, without the dosage level, the supplements cannot be determined to be a better source for tooth growth.